DOCUMENT RESUME

ED 204 795 CS 206 502

AUTHOR Flower, Linda: Hayes, John R.
TITLE Process-Based Evaluation of Writing: Changing the Performance, Not the Product.
PUB DATE Apr 81
NOTE 18p.: Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Discourse Analysis: Evaluation Methods: *Prewriting: Writing (Composition): *Writing Evaluation: Writing Processes: *Writing Research
IDENTIFIERS *Invention (Rhetorical): Protocol Analysis

ABSTRACT

This examination of an evaluation of writing based on process rather than on product argues that one of the primary functions of evaluation as a part of teaching should be to diagnose the writing strategies that underlie a writer's current performance, not just textual problems. The first half of the paper discusses the various ways teachers use evaluation: the purposes for their extensive, detailed editing and elaborate comment on student papers: and the reasons that such dedicated responses so often fail. It then reviews the research in the area of defining and influencing writers' thinking strategies. The second half of the paper describes a study of how writers generate ideas that contrasted the verbal strategies of four "good" writers with four "poor" writers as they composed aloud by analyzing transcripts of their verbal protocols with regard to new idea generation and new idea source. A discussion of the results of the study reports that good writers generated new ideas in response to a rhetorical problem, while the idea source for poor writers was either in response to the topic itself or to a current element in memory. (AEA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED204795

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PROCESS-BASED EVALUATION OF WRITING: CHANGING THE PERFORMANCE, NOT THE PRODUCT

Linda Flower and John R. Hayes

Carnegie-Mellon University

Pittsburgh, PA 15213

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Linda Flower

John R. Hayes

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

To appear in The Psychology of Composition. Ed. Douglas Buttruff. Conway, Arkansas: L&S Books, in press.

CS206502

Linda Flower
John R. Hayes
Carnegie-Mellon University
Pittsburgh, PA 15213

Process-Based Evaluation of Writing:

Changing the Performance, Not the Product

The psychology of writing is in large part the psychology of thinking, that is, the thinking processes people use when they write. By understanding those thinking processes we hope to teach them and, as a result, help writers produce better final products. But can this procedure work in reverse? Can we look at that final product and evaluate (or even guess at) the thinking processes that produced it? Let us put this question in practical terms: can teachers use papers to evaluate a student's writing process? Furthermore, can all those hours teachers spend on paper comments do more than change a given written product; can they actually change the writer's performance?

This is a paper about just such process-based evaluation of writing.[1] In it we will look at ways teachers commonly diagnose problems and at some new research on evaluating the writing process itself. Throughout, we will be arguing that one of the primary functions of evaluation as a part of teaching should be to diagnose not just textual problems but the writing strategies

---

[1] Much of the thinking in this paper originated in response to an NIE sponsored conference on Teachers' Response to Writing, September, 1978. The research was partially supported by a grant from the National Institute of Education, Department of Health, Education, and Welfare, Grant NIE G780195.

which underlie a writer's current performance.

In recent years research on evaluation has thrown itself into developing writing tests which meet the needs of educational administrators. We can, with tolerable certainty, predict whether a given student will or will not pass Freshman Composition in college. And we can say whether a group of writers, though not an individual, has improved or not after that year of Freshman Composition is over. Although this kind of summative evaluation is necessary, the immediate beneficiary is more likely to be an administrator who needs statistics than it is the student who can't write.

Teachers clearly use evaluation in different ways. In some cases a teacher's evaluation would be more accurately described as a "response" to writing in which the goal is not to teach writing per se but to increase the student's motivation through encouragement, to foster exploration or creativity, or simply to protect the tender ego writers are assumed to possess. In other words, evaluation is a tool, or rather a whole kit of tools, each designed for a particular purpose. Teachers usually rummage around in this kit, changing tools depending on what their goal is. For example, if the goal is to make an administrative decision, the teacher might use wholistic evaluation to compare groups of students. Later, in a conference, she might play muse and mentor, responding positively to anything promising in the text. Faced that evening with a pile of papers, she might then take the role of editor in which her goal would be to evaluate and improve this particular text through detailed comment and

atten*ion to errors.

In the role of editor, the teacher is primarily concerned with the student's current written product and the teacher's response or evaluation is designed to turn that product into an acceptable text. But what about the role of the teacher as teacher, when the goal is to change the student, not the current written product? Is editorial response the best method? And if not, are there other ways to diagnose and change the thinking process that underlies the student's writing problems? Let us take these two questions up in order.

Evaluating the Written Product

When teachers sincerely want to change a writer's performance, they often turn to extensive, detailed editing and elaborate comment (Brienza, 1979). Yet the limited research that has been done suggests that the hours spent marking comma splices in red and writing "awk org" in the margin may display heroism in teachers but have little effect on the subsequent production of errors by the student (Haynes, 1978; Wolter, 1976).

There must be a number of reasons such dedicated responses fail. One is that they leave much of the work we usually expect from teachers up to the student. They present the student with a rather random selection of individual errors--punctuation here, incoherence there--but all are deeply embedded in the context of a piece of writing. This means that the enormously difficult task of inferring the principle behind the instance is left up to the learner. From a small set of "disorganized" sentences, we

expect the learner to discover the hidden principle of good organization. Little wonder many fail to discover the magic principle, much less transfer it to a new writing task.

For example, when a reader or teacher writes "org" by the fourth sentence in a paragraph or says the paragraph is "disorganized," there is a temptation to see this as a local problem ("What's wrong with that sentence?") or to see it as a metaphysical problem ("It makes sense to me, so I guess I just can't express myself, or they just don't understand," etc.). Yet what the student may really need is a better principle for organizing the paragraph in the first place. For example, he or she may need to replace an impoverished principle, such as "Good organization is when everything follows when I read my paragraph over," with a more demanding principle, such as "A paragraph that seems clear when I read it to myself may not seem clear to a reader who doesn't know all that I do. A good organization which will work for my reader is one which makes the relations between ideas explicit right in the text. I can do this with a variety of tools, such as topic sentences, transitions, parallel structure, etc."

There is another reason detailed editing of the finished product can fail to teach someone how to write. Even if a student knows the principle behind, say "good organization," that abstract semantic knowledge may have little effect on their actual writing process. It may not help the student to recognize and change the thinking process which initially produced that incoherence and will probably produce it again. Here is a case in point. To many students, the hallmark of good writing is prose that "flows." In operational terms this means that if it

"flows" as you write it down, it is likely to be good (Sommers, 1978). When the goal is to produce prose which "flows" from one sentence to the next, a writer is likely to concentrate on creating logical or even merely semantic connections between adjacent sentences, but to ignore the larger logical or rhetorical focus which should hold that paragraph together. Furthermore, when writers reread what they have written, they may judge that it passes their flow test and feel satisfied, because they find good continuity between adjacent sentences. Unfortunately this composing and testing strategy, which is a very local one, may in fact direct the writer's attention away from matters of general organization. In this case, poor organization is not due to a lack of knowledge per se, but may be the result of a very positively undertaken writing stragegy.

Evaluating the Writer's Strategies

We think evaluation can play a key role in teaching people to write when it is used to diagnose the strategies writers use in composing and to help them refine or alter those strategies. One model for this approach is the work of Albert Stevens and Allan Collins who studied the teaching techniques of experienced Socratic tutors. They were able first to define a set of twenty-four specific strategies tutors use, such as asking for a prediction or giving a counter-example (Collins, 1977). Secondly, they showed how tutors used these teaching strategies not simply to correct individual faults, but to refine the student's own causal model of some body of ideas, such as their

model of how rising air is related to rainfall. Then the tutors tried to refine the student's ability to use that model (Stevens and Collins, 1977). As Stevens and Collins showed, the diagnosis and correction made by these tutors was driven by a higher goal which was to change the student's thinking strategies, not his or her current answer.

This use of evaluation, to diagnose process, is certainly not new to teachers: Stevens and Collins' attempt to formalize the tutoring procedure and develop a Computer Aided Instruction system was based on studying successful teachers.

Likewise, much of the enormous impact of Mina Shaughnessy's Errors and Expectations was due to the fact that she redefined the errors of Basic Writers into understandable verbal patterns and speech strategies. By revealing a logic behind what appeared to be a bewildering mass of error, she helped students and teachers see a principle that governs a wide range of practice. With a similar goal, one of the present authors has described the phenomenon of Writer-based prose as a common writing problem with an underlying process-based logic. Although the situation demands Reader-based prose, organized with the reader's needs in mind, people resort to the private monologues of Writer-based prose for sound psychological reasons--at least from the writer's point of view. In Writer-based prose, the writer's egocentric focus and use of a narrative organization or simple survey of his or her own knowledge are highly effective strategies for retrieving ideas and producing flowing prose. Unfortunately, this composing strategy often fails to meet the reader's needs

for reader-based, hierarchically structured discussion.

One of the goals of research then is to discover some of the principles that govern practice, and to help teachers evaluate and alter the strategies that poor writers use. In the remainder of this paper we will describe one such study focused on two parallel strategies good and poor writers use for the task of generating ideas.

Evaluating How Writers Generate Ideas

This study started with the hypothesis that an important difference between good and poor writers lies, not simply in their ability to express ideas in written speech, but in the very strategies they use to generate those ideas in the first place. We had observed that poor writers, in their attempt to find a focus or thesis for a paper, often seemed tied to the topic, while more experienced writers appeared to be responding to a larger rhetorical problem--a problem which included the reader and their own goals (Flower and Hayes, in press).

If this hypothesis were true it would mean a number of things. First, if poor writers are obsessively focused on the topic to the exclusion of the larger rhetorical problem, it could help explain why they often are more likely to violate conventions of the appropriate when they write, but not when they speak. In face-to-face conversation only ancient mariners are likely to ignore the rhetorical situation. Secondly, if the rhetorical situation itself helps stimulate plans and ideas, it could explain why so many poor writers, including the ones in our

experiment, often seem to "run out of ideas." Finally, if a significant difference between good and poor writers is the strategy they use to generate ideas, this would suggest that evaluating or editing the final product qua product is unlikely to produce dramatic change. A more effective teaching technique would focus on the writing process itself.

To test this hypothesis we collected, on tape, verbal protocols of nine writers composing aloud. Four were people on the university staff who liked to write, who had done writing, and who were considered by their peers to be "good" writers. In contrast to this group of "good" or experienced writers, we studied a group of "poor" writers who had come or been sent to the Communications Skills Center for general problems with organization and coherence. Two of these "poor" writers were graduate students, two undergraduate; none had difficulty with basic grammar or sentence structure.

Each subject wrote on an assignment about which they would have topic information, but which created significant audience constraints. They were told to work for approximately an hour and to verbalize everything that went through their minds as they wrote. We analyzed these transcripts in two steps: first, by isolating each new idea that was generated during the session (see Table 1). A new idea was defined as any complete grammatical unit, including complex statements with dependent clauses. However, if such a complex statement was generated in two attempts separated by long pauses or intervening material, it was coded as two ideas. Our goal was to code as one unit those

ideas which were being retrieved from memory as a unit, and to code as new all new attempts to expand or develop an idea. Changes which affected merely the wording or sentence structure were judged on the basis of our model of the writing process to belong to the process of "translating," not "generating," and were not counted (Hayes and Flower, 1979).

Insert Table 1 here

Our second step was to discover where each of these new ideas came from. Therefore we categorized each new idea in one of three ways, as a response to either:

1. the larger rhetorical problem, including the topic
2. the topic alone
3. a current element in memory

The basic purpose of this categorization was simple. Within a given body of ideas, we wanted to see what per cent of those ideas were generated as a response merely to the topic alone (or to a current element in memory) or alternatively, as a response to not only the topic, but to the larger rhetorical situation as well.

We used the following taxonomy to decide into which category a new idea should be placed:

1. An idea was categorized as a response to the rhetorical problem if it indicated one of the following: a concern with the writer's Purpose or Goal; an indication of the writer's sense of

Audience; or a concern with the writer's sense of Self or Persona. A writer's concern with purpose or Goals took two major forms: either as a statement of purpose (e.g., "I need something here that pulls you in") or as a recognition of some of the formal features of written text. Statements such as, "I'll use this as an introduction," indicated that the writer was seeing her ideas in the larger context of writing a paper or article. Ideas which responded to the audience were sometimes direct ("Make this friendlier; it's for a young audience") or implicit (e.g., "I'll list the names of the most fascinating drinks"). Writers generated ideas in response to their projected Self or Persona with comments such as, "I'll appear like an idiot." Finally, some statements combined a number of these elements, such as "This may not be the best term for ten-year-olds, but it maintains the rhythm." Any idea which showed some response to the larger rhetorical problem, then, was placed in the first category.

2. The second category was reserved for new ideas which were judged as simply information generated in response to the topic alone, such as "A waitress has a number of duties, first . . . ." Often these ideas appeared to be the result of a straight memory search of what the writer remembered about the topic.

3. The final category was necessary to account for new ideas which appeared to be connected by some association to a recent thought or current item in memory, but which were not relevant to the rhetorical problem or to the topic. Both good and poor writers appear to go off on these short trains of association;

the difference is in how frequently they do it.

As we expected, it is difficult to get complete agreement from judges on the absolute number of new ideas, since different judges are likely to have different thresholds for distinguishing new ideas from mere rewordings of old ones. Therefore, as a check on our method we conducted the following test using a set of 73 ideas which were judged as new ideas by all four judges in the experiment. The critical judgments in this study are the categorizations: is a given idea a response to the rhetorical problem or merely to the topic or a current element? And can judges agree on making those categorizations? In a universe of 73 ideas which all four judges had selected as new ideas, there was complete categorization agreement on 52 ideas or 71%. Three judges agreed on 69 responses or 94.5%. There was a two or three-way split on only four ideas. This percentage of agreement confirmed our sense that such categorization is reliable.

Results

The results of this analysis were striking, especially since our subjects did not represent the extremes of either good or poor writers and there was no attempt to account for or control individual differences. Nevertheless, as Table 1 shows, the poor writers as a group generated only 28% of their new ideas in response to the rhetorical problem; the other 72% were in response to the topic and/or a current element in memory. For the good writers this 30/70 distribution was nearly reversed. Good writers generated 60% of their new ideas in response to the

rhetorical problem in some way; only 40% of their ideas were a response to th topic or current element alone.

As you can see in Table 1, there is a significant spread among the poor writers, but as a group they remained distinct from the good writers. On the basis of a Pitman randomization test, these results are significant at the .01 level; that is, the probability that this difference between the groups would have arisen by chance is less than 1 in 100 (Siegel, 1956).

This study has, we hope, implications for teaching in two ways: first, as a study of a particular composing problem and secondly, as an example of the kind of research that will enable us to offer process-based evaluation to students. As a study of how writers generate ideas, this project and others support the growing attempt to teach the art of invention as a part of the writing process. But it does so with a difference. Getting writers to "freewrite," to roam through memory dredging up ideas on the topic, may not be enough. One reason poor writers may have trouble producing ideas is that they are not taking advantage of all the cues or stimuli to new ideas offered by the rhetorical problem. Thinking about one's Purpose, Audience, or Persona is a powerful technique for discovering topic-based things to say. Furthermore, teachers can have a significant impact on how much of the rhetorical problem students do consider. As Scardamalia, Bracewell, and Bereiter have shown, the number of audience cues teachers give in an assignment has a direct effect on the student's ability to decenter and consider his or her audience (Scardamalia, 1978).

How does this study fit into the larger question of evaluation? There is no doubt that there are many things which contribute to the performance of any given poor writer. However, studies such as these will help us find those strategies which make a difference. The sort of detailed protocol analysis used here is clearly a research tool; it is too complicated and time consuming to use in evaluating individual writers. However, it can contribute to teaching by defining significant process-based problems. This in turn allows teachers to diagnose the process behind their students' performance and to design assignments and instruction which help writers abandon unproductive writing strategies for ones that work. Because such diagnosis focuses on the writer at work, it lets us intervene where we might do the most good--by changing the writer's thinking process, not just the final product.

References

Bracewell, Robert J., Marlene Scardamalia and Carl Bereiter. "Writing and Decentered Thought: The Development of Audience Awareness." American Educational Research Association meeting, Toronto, March 27-31, 1978 (ERIC Ed 154 433)

Brienza, Susan D. "The Style of Teacher Response." Conference on English Education, Symposium on Cognition and Written Language, Pittsburgh, PA, March 15, 1979.

Collins, Allan. "Processes in Acquiring Knowledge." In Schooling and the Acquisition of Knowledge. Eds. R. C. Anderson, R. J. Spiro, W. E. Montague. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1977.

Flower, Linda, and John R. Hayes. "The Cognition of Discovery: Defining A Rhetorical Problem." College Composition and Communication, in press.

Flower, Linda. "Writer-Based Prose: A Cognitive Basis for Problems in Writing." College English, 41 (September, 1979), 19-37.

Hayes, John R., and Linda Flower. "Identifying the Organization of Writing Processes." In Cognitive Processes in Writing. Eds. Lee Gregg and Erwin Steinberg. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1979

Haynes, Elizabeth. "Using Research in Preparing to Teach Writing." English Journal, LXVII (Jan., 1978), pp. 82-88.

Shaughnessy, Mina. Errors and Expectations. New York: Oxford, 1977.

Siegel, S. Nonparametric Methods for Behavioral Sciences. New York: McGraw-Hill, 1956.

Sommers, Nancy I. "Revision Strategies of Student Writers and Experienced Writers." College Composition and Communication, in press.

Stevens, Albert, and Allan Collins. "The Goal Structure of a Socratic Tutor." Report No. 3518, Bolt Beranek and Newman Inc., 50 Moulton St., Cambridge, Mass. 02138, March, 1977.

Wolter, Daniel, and Walter Lamberg. "Research on the Effect of of Feedback on Writing: Review and Implications." Austin, TX: University of Texas, 1976 (ERIC No. 140 355).

		Total New Ideas	Response to Rhetorical Problem	Response to Topic	Response to Current Element	% Response Rhetorical Problem	% Response to Topic & Current Element	Group Ratio
Good Writers	G1	158	106	48	4	67%	33%	
	G2	93	56	30	7	60	40	
	G3	115	66	45	4	57	42	
	G4	85	45	36	4	53	47	
Average		113	68	40	5	61%	39%	61/39
Poor Writers	P1	129	22	66	41	17%	83%	
	P2	85	20	63	2	24	76	
	P3	53	15	37	1	28	72	
	P4	134	41	77	16	31	69	
	P5	67	31	36	0	46	54	
Average		94	26	56	12	28%	72%	28/72

Table 1 Differences in How Good and Poor Writers Generate New Ideas